Multifrequency, Multiresolution Image Detection

Andrew J. Connolly
Department of Physics and Astronomy
University of Pittsburgh
3941 O'Hara Street (Allen Hall)
Pittsburgh PA 15260

NASA Grant AISR NAG5 9399: May 2000 - May 2003

Original Abstract of Three-Year NASA AISR Program

The next decade will mark the start of a new era in multi-frequency astronomy. New surveys and satellites are coming on line that will open up the full electromagnetic spectra at higher precision and to fainter flux levels than ever before (from X-rays through to the submillimeter and radio). Together these multi-frequency surveys will provide a bolometric view of the Universe from which we might fully reconstruct the properties of Galactic and extra-galactic sources. Studying these data sets in isolation will provide some insight into the processes that drive the evolution of our Universe. It is only, however, if we analyze the multi-frequency data in unison that we might fully exploit their scientific potential. We have, therefore, developed a new and novel technique for combining multi-frequency data that provides objective source detection. We propose here to develop these algorithms into a suite of software tools that will enable the astronomical community to explore multi-frequency, multi-resolution data in an optimal fashion.

Goals and comparison with accomplishments/findings

The goals of this AISR program were to develop algorithms to analyze multifrequency and multispectral data. As part of this research we focused on two components: statistics and algorithms for defining thresholds within multifrequency data and the extension of source detection algorithms to account for multispectral data. The latter of these components focused on the integration of source detection with multispectral catalogs of data.

The first year of this proposal saw the development and testing of algorithms used in multifrequency image detection and analysis. This was accomplished by constructing a prototype pipeline and simulation software on which these algorithms could be tested. Two prototype pipelines based on PERL and IDL were developed and tested. These systems input multifrequency images (in FITS format) and then flatten and regularized the images (accounting for the presence of bright sources across an image which affect the background and noise properties). All multifrequency images were combined to form a χ^2 image and then the SExtractor image detection package was applied to this probability image to identify sources.

This initial pipeline demonstrated that one of the most important constraints on a detection algorithm is the detection threshold applied to the data. As we decrease the threshold we detect more sources at the cost of additional spurious detections. We therefore developed a technique (False Discovery Rate, Miller et al 2001 and Hopkins et al 2001) that enables the number of false detections to be controlled. This approach was found to be well suited to the detection of faint sources as it *a priori* specifies the acceptable number of contaminants within any sample of galaxies (i.e. false positives). The number of spurious sources is, as expected, better controlled with this approach than with the earlier Bayesian thresholding. We also found that this new thresholding is relatively robust to the point-spread-function of the telescope. This has the potential of making our approach directly applicable to any set of imaging data (from X-ray through to infrared and radio wavelengths).

The second year focused on the analysis of multifrequency NASA data. Using high resolution images from local and distant multicolor imaging surveys (e.g. the Hubble Deep Fields and the Sloan Digital Sky Survey) we determined the spectral energy distributions of the individual pixels within an image (by fitting a range of template spectra that represent different star formation histories, extinctions etc). Once the spectral energy distribution of an individual pixel are known we can predict the color of that pixel as a function of filter response function or redshift. By specifying the telescope design parameters (e.g. pixel scale and angular resolution) we can relate sources from taken with different observing systems (Conti et al 2003).

The third year focused on integrating source extraction techniques with the Virtual Observatory to enable the cross-correlation of multispectral data sets. Our initial application (WESIX; web enabled source extraction and cross-matching) takes an input image, runs SExtractor to identify sources, cross-matches these sources with other datasets available within the VO and delivers to a user catalogs with matched sources (as well as sources that are unmatched in either the input or VO catalog). Sources can be visualized through the Aladin image display plugin and the correlation between the input image data and the cross-matched catalog can be viewed using VOPlot. This approach not only provides a simple source detection it enables automated photometric and astrometric calibration of images. WESIX integrates source extraction and the VO through the openskyquery protocol and is now available through the University of Pittsburgh's NVO portal (http://nvogre.phyast.pitt.edu/gestalt)

Papers resulting from this proposal

"The Star Formation History of Galaxies Measured from Individual Pixels. I. The Hubble Deep Field North", Conti, A., Connolly, A.J., Hopkins, A., Budavari, T., Csabai, I., Szalay, A.S., Schmidt, S., Adams, C., and Petrovic, N., 2003, AJ, 126, 2330
"Controlling the False Discovery Rate in Astrophysical Data Analysis", Miller, Genovese, Nichol, Wasserman, Connolly, Reichart, Hopkins, Schneider, Moore, 2001, AJ, 122, 3492
"A New Source Detection Algorithm using FDR", Hopkins, Miller, Connolly, Genovese, Nichol, Wasseramn, 2001 AJ, 123, 1086